BIOTECHNOLOGY
SYSTEMS·
BRANCH

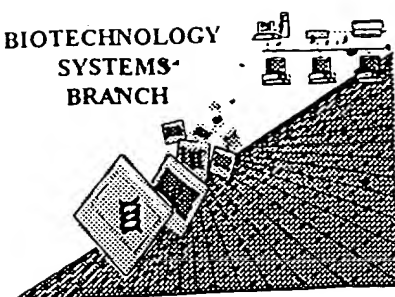# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/761534_

Source: _OIPE_

Date Processed by STIC: _10/11/01_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

```
                 RAW SEQUENCE LISTING              DATE: 10/11/2001
                 PATENT APPLICATION:  US/09/761,534    TIME: 13:34:44

                 Input Set : A:\0399.2006-003SEQLIST.txt
                 Output Set: N:\CRF3\10112001\I761534.raw

   4 <110> APPLICANT: Huang, Qian
   5        Richmond, Joan F.L.
   6        Cho, Bryan K.
   7        Palliser, Deborah
   8        Chen, Jianzhu
   9        Eisen, Herman N.
  10        Young, Richard A.
  12 <120> TITLE OF INVENTION: In Vivo CTL Elicitation By Heat Shock
  13        Protein Fusion Proteins Maps To A Discrete Domain and is
  14        CD4+T Cell-Independent
  17 <130> FILE REFERENCE: 0399.2006-003
  19 <140> CURRENT APPLICATION NUMBER: US 09/761,534
  20 <141> CURRENT FILING DATE: 2001-01-16
  22 <150> PRIOR APPLICATION NUMBER: PCT/US00/32831
  23 <151> PRIOR FILING DATE: 2000-12-01
  25 <150> PRIOR APPLICATION NUMBER: US 60/176,143
  26 <151> PRIOR FILING DATE: 2000-01-14
  28 <160> NUMBER OF SEQ ID NOS: 25
  30 <170> SOFTWARE: FastSEQ for Windows Version 4.0
  32 <210> SEQ ID NO: 1
  33 <211> LENGTH: 8
  34 <212> TYPE: PRT
  35 <213> ORGANISM: Unknown
  37 <220> FEATURE:
  38 <223> OTHER INFORMATION: Peptide Liberated From P1
  40 <400> SEQUENCE: 1
  41 Ser Ile Tyr Arg Tyr Tyr Gly Leu
  42  1               5
  45 <210> SEQ ID NO: 2
  46 <211> LENGTH: 8
  47 <212> TYPE: PRT
  48 <213> ORGANISM: Unknown
  50 <220> FEATURE:
  51 <223> OTHER INFORMATION: Ova Peptide
  53 <400> SEQUENCE: 2
  54 Ser Ile Ile Asn Phe Glu Lys Leu
  55  1               5
  58 <210> SEQ ID NO: 3
  59 <211> LENGTH: 8
  60 <212> TYPE: PRT
  61 <213> ORGANISM: Unknown
  63 <220> FEATURE:
  64 <223> OTHER INFORMATION: Alpha KG Peptide
  66 <400> SEQUENCE: 3
  67 Leu Ser Pro Phe Pro Phe Asp Leu
  68  1               5
  71 <210> SEQ ID NO: 4
```

Does Not Comply
Corrected Diskette Needed

Does Not Comply
Corrected Diskette Needed

220 and 223 Artificial Sequence

check diskette

RAW SEQUENCE LISTING   DATE: 10/11/2001
PATENT APPLICATION: US/09/761,534   TIME: 13:34:44

Input Set : A:\0399.2006-003SEQLIST.txt
Output Set: N:\CRF3\10112001\I761534.raw

```
 72 <211> LENGTH: 5
 73 <212> TYPE: PRT
 74 <213> ORGANISM: Unknown
 76 <220> FEATURE:
 77 <223> OTHER INFORMATION: Octapeptide
 79 <400> SEQUENCE: 4
 80 Ser Tyr Arg Gly Leu
 81 1                5
 84 <210> SEQ ID NO: 5
 85 <211> LENGTH: 1260
 86 <212> TYPE: DNA
 87 <213> ORGANISM: Mycobacterium Tuberculosis hsp70 cDNA
 89 <220> FEATURE:
 90 <221> NAME/KEY: CDS
 91 <222> LOCATION: (0)...(1260)
 93 <400> SEQUENCE: 5
 94 atg gct cgt gcg gtc ggg atc gac ctc ggg acc acc aac tcc gtc gtc    48
 95 Met Ala Arg Ala Val Gly Ile Asp Leu Gly Thr Thr Asn Ser Val Val
 96 1               5                   10                  15
 98 tcg gtt ctg gaa ggt ggc gac ccg gtc gtc gtc gcc aac tcc gag ggc    96
 99 Ser Val Leu Glu Gly Gly Asp Pro Val Val Val Ala Asn Ser Glu Gly
100                 20                  25                  30
102 tcc agg acc acc ccg tca att gtc gcg ttc gcc cgc aac ggt gag gtg   144
103 Ser Arg Thr Thr Pro Ser Ile Val Ala Phe Ala Arg Asn Gly Glu Val
104             35                  40                  45
106 ctg gtc ggc cag ccc gcc aag aac cag gca gtg acc aac gtc gat cgc   192
107 Leu Val Gly Gln Pro Ala Lys Asn Gln Ala Val Thr Asn Val Asp Arg
108         50                  55                  60
110 acc gtg cgc tcg gtc aag cga cac atg ggc agc gac tgg tcc ata gag   240
111 Thr Val Arg Ser Val Lys Arg His Met Gly Ser Asp Trp Ser Ile Glu
112 65                  70                  75                  80
114 att gac ggc aag aaa tac acc gcg ccg gag atc agc gcc cgc att ctg   288
115 Ile Asp Gly Lys Lys Tyr Thr Ala Pro Glu Ile Ser Ala Arg Ile Leu
116                 85                  90                  95
118 atg aag ctg aag cgc gac gcc gag gcc tac ctc ggt gag gac att acc   336
119 Met Lys Leu Lys Arg Asp Ala Glu Ala Tyr Leu Gly Glu Asp Ile Thr
120                 100                 105                 110
122 gac gcg gtt atc acg acg ccc gcc tac ttc aat gac gcc cag cgt cag   384
123 Asp Ala Val Ile Thr Thr Pro Ala Tyr Phe Asn Asp Ala Gln Arg Gln
124         115                 120                 125
126 gcc acc aag gac gcc ggc cag atc gcc ggc ctc aac gtg ctg cgg atc   432
127 Ala Thr Lys Asp Ala Gly Gln Ile Ala Gly Leu Asn Val Leu Arg Ile
128     130                 135                 140
130 gtc aac gag ccg acc gcg gcc gcg ctg gcc tac ggc ctc gac aag ggc   480
131 Val Asn Glu Pro Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Gly
132 145                 150                 155                 160
134 gag aag gag cag cga atc ctg gtc ttc gac ttg ggt ggt ggc act ttc   528
135 Glu Lys Glu Gln Arg Ile Leu Val Phe Asp Leu Gly Gly Gly Thr Phe
136                 165                 170                 175
```

RAW SEQUENCE LISTING   DATE: 10/11/2001
PATENT APPLICATION: US/09/761,534   TIME: 13:34:44

Input Set : A:\0399.2006-003SEQLIST.txt
Output Set: N:\CRF3\10112001\I761534.raw

```
138 gac gtt tcc ctg ctg gag atc ggc gag ggt gtg gtt gag gtc cgt gcc   576
139 Asp Val Ser Leu Leu Glu Ile Gly Glu Gly Val Val Glu Val Arg Ala
140             180                 185                 190
142 act tcg ggt gac aac cac ctc ggc ggc gac gac tgg gac cag cgg gtc   624
143 Thr Ser Gly Asp Asn His Leu Gly Gly Asp Asp Trp Asp Gln Arg Val
144             195                 200                 205
146 gtc gat tgg ctg gtg gac aag ttc aag ggc acc agc ggc atg gat ctg   672
147 Val Asp Trp Leu Val Asp Lys Phe Lys Gly Thr Ser Gly Met Asp Leu
148         210                 215                 220
150 acc aag gac aag atg gcg atg cag cgg ctg cgg gaa gcc gcc gag aag   720
151 Thr Lys Asp Lys Met Ala Met Gln Arg Leu Arg Glu Ala Ala Glu Lys
152 225                 230                 235                 240
154 gca aag atc gag ctg agt tcg agt cag tcc acc tcg atc aac ctg ccc   768
155 Ala Lys Ile Glu Leu Ser Ser Ser Gln Ser Thr Ser Ile Asn Leu Pro
156             245                 250                 255
158 tac atc acc gtc gac gcc gac aag aac ccg ttg ttc tta gac gag cag   816
159 Tyr Ile Thr Val Asp Ala Asp Lys Asn Pro Leu Phe Leu Asp Glu Gln
160             260                 265                 270
162 ctg acc cgc gcg gag ttc caa cgg atc act cag gac ctg ctg gac cgc   864
163 Leu Thr Arg Ala Glu Phe Gln Arg Ile Thr Gln Asp Leu Leu Asp Arg
164         275                 280                 285
166 act cgc aag ccg ttc cag tcg gtg atc gct gac acc ggc att tcg gtg   912
167 Thr Arg Lys Pro Phe Gln Ser Val Ile Ala Asp Thr Gly Ile Ser Val
168     290                 295                 300
170 tcg gag atc gat cac gtt gtg ctc gtg ggt ggt tcg acc cgg atg ccc   960
171 Ser Glu Ile Asp His Val Val Leu Val Gly Gly Ser Thr Arg Met Pro
172 305                 310                 315                 320
174 gcg gtg acc gat ctg gtc aag gaa ctc acc ggc ggc aag gaa ccc aac  1008
175 Ala Val Thr Asp Leu Val Lys Glu Leu Thr Gly Gly Lys Glu Pro Asn
176             325                 330                 335
178 aag ggc gtc aac ccc gat gag gtt gtc gcg gtg gga gcc gct ctg cag  1056
179 Lys Gly Val Asn Pro Asp Glu Val Val Ala Val Gly Ala Ala Leu Gln
180             340                 345                 350
182 gcc ggc gtc ctc aag ggc gag gtg aaa gac gtt ctg ctg ctt gat gtt  1104
183 Ala Gly Val Leu Lys Gly Glu Val Lys Asp Val Leu Leu Leu Asp Val
184         355                 360                 365
186 acc ccg ctg agc ctg ggt atc gag acc aag ggc ggg gtg atg acc agg  1152
187 Thr Pro Leu Ser Leu Gly Ile Glu Thr Lys Gly Gly Val Met Thr Arg
188     370                 375                 380
190 ctc atc gag cgc aac acc acg atc ccc acc aag cgg tcg gag act ttc  1200
191 Leu Ile Glu Arg Asn Thr Thr Ile Pro Thr Lys Arg Ser Glu Thr Phe
192 385                 390                 395                 400
194 acc acc gcc gac gac aac caa ccg tcg gtg cag atc cag gtc tat cag  1248
195 Thr Thr Ala Asp Asp Asn Gln Pro Ser Val Gln Ile Gln Val Tyr Gln
196             405                 410                 415
198 ggg gag cgt gag                                                   1260
199 Gly Glu Arg Glu
200             420
203 <210> SEQ ID NO: 6
```

```
204 <211> LENGTH: 420
205 <212> TYPE: PRT
206 <213> ORGANISM: Mycobacterium Tuberculosis hsp70 cDNA
208 <400> SEQUENCE: 6
209 Met Ala Arg Ala Val Gly Ile Asp Leu Gly Thr Thr Asn Ser Val Val
210  1               5                   10                  15
211 Ser Val Leu Glu Gly Gly Asp Pro Val Val Val Ala Asn Ser Glu Gly
212              20                  25                  30
213 Ser Arg Thr Thr Pro Ser Ile Val Ala Phe Ala Arg Asn Gly Glu Val
214              35                  40                  45
215 Leu Val Gly Gln Pro Ala Lys Asn Gln Ala Val Thr Asn Val Asp Arg
216      50                  55                  60
217 Thr Val Arg Ser Val Lys Arg His Met Gly Ser Asp Trp Ser Ile Glu
218 65                  70                  75                  80
219 Ile Asp Gly Lys Lys Tyr Thr Ala Pro Glu Ile Ser Ala Arg Ile Leu
220              85                  90                  95
221 Met Lys Leu Lys Arg Asp Ala Glu Ala Tyr Leu Gly Glu Asp Ile Thr
222              100                 105                 110
223 Asp Ala Val Ile Thr Thr Pro Ala Tyr Phe Asn Asp Ala Gln Arg Gln
224              115                 120                 125
225 Ala Thr Lys Asp Ala Gly Gln Ile Ala Gly Leu Asn Val Leu Arg Ile
226      130                 135                 140
227 Val Asn Glu Pro Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Gly
228 145                 150                 155                 160
229 Glu Lys Glu Gln Arg Ile Leu Val Phe Asp Leu Gly Gly Gly Thr Phe
230              165                 170                 175
231 Asp Val Ser Leu Leu Glu Ile Gly Glu Gly Val Val Glu Val Arg Ala
232              180                 185                 190
233 Thr Ser Gly Asp Asn His Leu Gly Gly Asp Asp Trp Asp Gln Arg Val
234              195                 200                 205
235 Val Asp Trp Leu Val Asp Lys Phe Lys Gly Thr Ser Gly Met Asp Leu
236      210                 215                 220
237 Thr Lys Asp Lys Met Ala Met Gln Arg Leu Arg Glu Ala Ala Glu Lys
238 225                 230                 235                 240
239 Ala Lys Ile Glu Leu Ser Ser Ser Gln Ser Thr Ser Ile Asn Leu Pro
240              245                 250                 255
241 Tyr Ile Thr Val Asp Ala Asp Lys Asn Pro Leu Phe Leu Asp Glu Gln
242              260                 265                 270
243 Leu Thr Arg Ala Glu Phe Gln Arg Ile Thr Gln Asp Leu Leu Asp Arg
244              275                 280                 285
245 Thr Arg Lys Pro Phe Gln Ser Val Ile Ala Asp Thr Gly Ile Ser Val
246      290                 295                 300
247 Ser Glu Ile Asp His Val Val Leu Val Gly Gly Ser Thr Arg Met Pro
248 305                 310                 315                 320
249 Ala Val Thr Asp Leu Val Lys Glu Leu Thr Gly Gly Lys Glu Pro Asn
250              325                 330                 335
251 Lys Gly Val Asn Pro Asp Glu Val Val Ala Val Gly Ala Ala Leu Gln
252              340                 345                 350
253 Ala Gly Val Leu Lys Gly Glu Val Lys Asp Val Leu Leu Leu Asp Val
```

RAW SEQUENCE LISTING   DATE: 10/11/2001
PATENT APPLICATION: US/09/761,534   TIME: 13:34:44

                Input Set : A:\0399.2006-003SEQLIST.txt
                Output Set: N:\CRF3\10112001\I761534.raw

254   355   360   365
255 Thr Pro Leu Ser Leu Gly Ile Glu Thr Lys Gly Gly Val Met Thr Arg
256   370   375   380
257 Leu Ile Glu Arg Asn Thr Thr Ile Pro Thr Lys Arg Ser Glu Thr Phe
258 385   390   395   400
259 Thr Thr Ala Asp Asp Asn Gln Pro Ser Val Gln Ile Gln Val Tyr Gln
260   405   410   415
261 Gly Glu Arg Glu
262   420
265 <210> SEQ ID NO: 7
266 <211> LENGTH: 630
267 <212> TYPE: DNA
268 <213> ORGANISM: Unknown
270 <220> FEATURE:
271 <223> OTHER INFORMATION: Segment II of TBhsp70αιℓ
273 <221> NAME/KEY: CDS
W--> 274 <222> LOCATION: (0)...(631)
275 <223> OTHER INFORMATION: Segment II of TBhsp70 αℓℓ
277 <400> SEQUENCE: 7
278 gag aag gag cag cga atc ctg gtc ttc gac ttg ggt ggt ggc act ttc   48
279 Glu Lys Glu Gln Arg Ile Leu Val Phe Asp Leu Gly Gly Gly Thr Phe
280 1   5   10   15
282 gac gtt tcc ctg ctg gag atc ggc gag ggt gtg gtt gag gtc cgt gcc   96
283 Asp Val Ser Leu Leu Glu Ile Gly Glu Gly Val Val Glu Val Arg Ala
284   20   25   30
286 act tcg ggt gac aac cac ctc ggc ggc gac gac tgg gac cag cgg gtc   144
287 Thr Ser Gly Asp Asn His Leu Gly Gly Asp Asp Trp Asp Gln Arg Val
288   35   40   45
290 gtc gat tgg ctg gtg gac aag ttc aag ggc acc agc ggc atg gat ctg   192
291 Val Asp Trp Leu Val Asp Lys Phe Lys Gly Thr Ser Gly Met Asp Leu
292   50   55   60
294 acc aag gac aag atg gcg atg cag cgg ctg cgg gaa gcc gcc gag aag   240
295 Thr Lys Asp Lys Met Ala Met Gln Arg Leu Arg Glu Ala Ala Glu Lys
296 65   70   75   80
298 gca aag atc gag ctg agt tcg agt cag tcc acc tcg atc aac ctg ccc   288
299 Ala Lys Ile Glu Leu Ser Ser Ser Gln Ser Thr Ser Ile Asn Leu Pro
300   85   90   95
302 tac atc acc gtc gac gcc gac aag aac ccg ttg ttc tta gac gag cag   336
303 Tyr Ile Thr Val Asp Ala Asp Lys Asn Pro Leu Phe Leu Asp Glu Gln
304   100   105   110
306 ctg acc cgc gcg gag ttc caa cgg atc act cag gac ctg ctg gac cgc   384
307 Leu Thr Arg Ala Glu Phe Gln Arg Ile Thr Gln Asp Leu Leu Asp Arg
308   115   120   125
310 act cgc aag ccg ttc cag tcg gtg atc gct gac acc ggc att tcg gtg   432
311 Thr Arg Lys Pro Phe Gln Ser Val Ile Ala Asp Thr Gly Ile Ser Val
312   130   135   140
314 tcg gag atc gat cac gtt gtg ctc gtg ggt ggt tcg acc cgg atg ccc   480
315 Ser Glu Ile Asp His Val Val Leu Val Gly Gly Ser Thr Arg Met Pro
316 145   150   155   160

```
<210> SEQ ID NO 16
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION:
<400> SEQUENCE: 16
      ggaattccta tctagtcact tgccctcccg gccgtc
36
```

Errored

actual file contents
as of 11/05/01
2:03 pm

A 213 response of "unknown" requires
an explanation in field 223. MA

VERIFICATION SUMMARY   DATE: 10/11/2001
PATENT APPLICATION: US/09/761,534   TIME: 13:34:45

Input Set : A:\0399.2006-003SEQLIST.txt
Output Set: N:\CRF3\10112001\I761534.raw

L:91 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:5, CDS LOCATION: (0)...
(1260)
L:274 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:7, CDS LOCATION: (0)...
(631)
L:383 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:9, CDS LOCATION: (0)...
(1929)
L:654 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:11, CDS LOCATION: (0)...
(627)
L:793 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:793 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:

Errored